Subgraph Epimorphisms: Theory and Application to Model Reductions in Systems Biology

François Fages
Joint work with Steven Gay (PhD Thesis) and Sylvain Soliman

LIFEWARE Team
Inria Paris-Rocquencourt

CIRM Marseille, May 28, 2015

Systems Biology

Biologists use diagrammatic notations. Kohn's map of cell cycle:

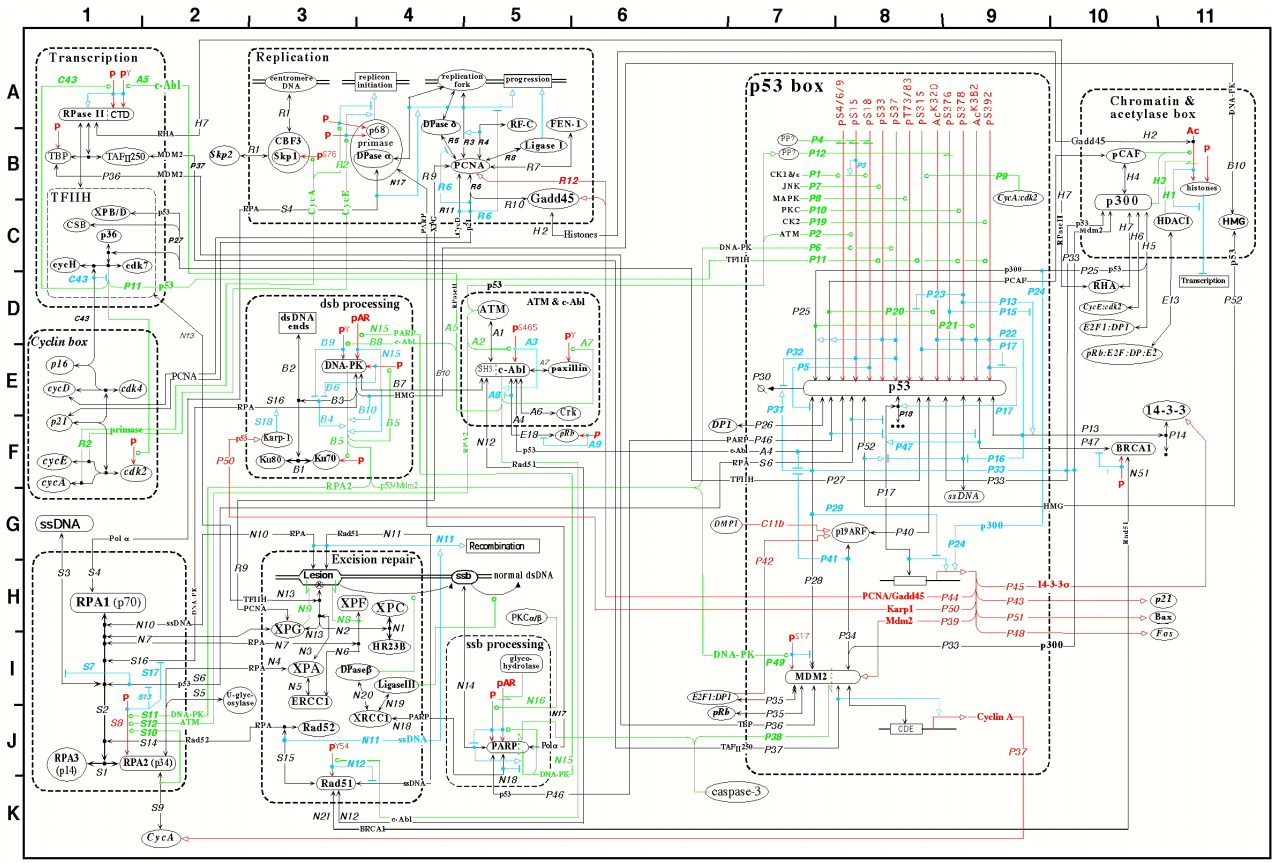
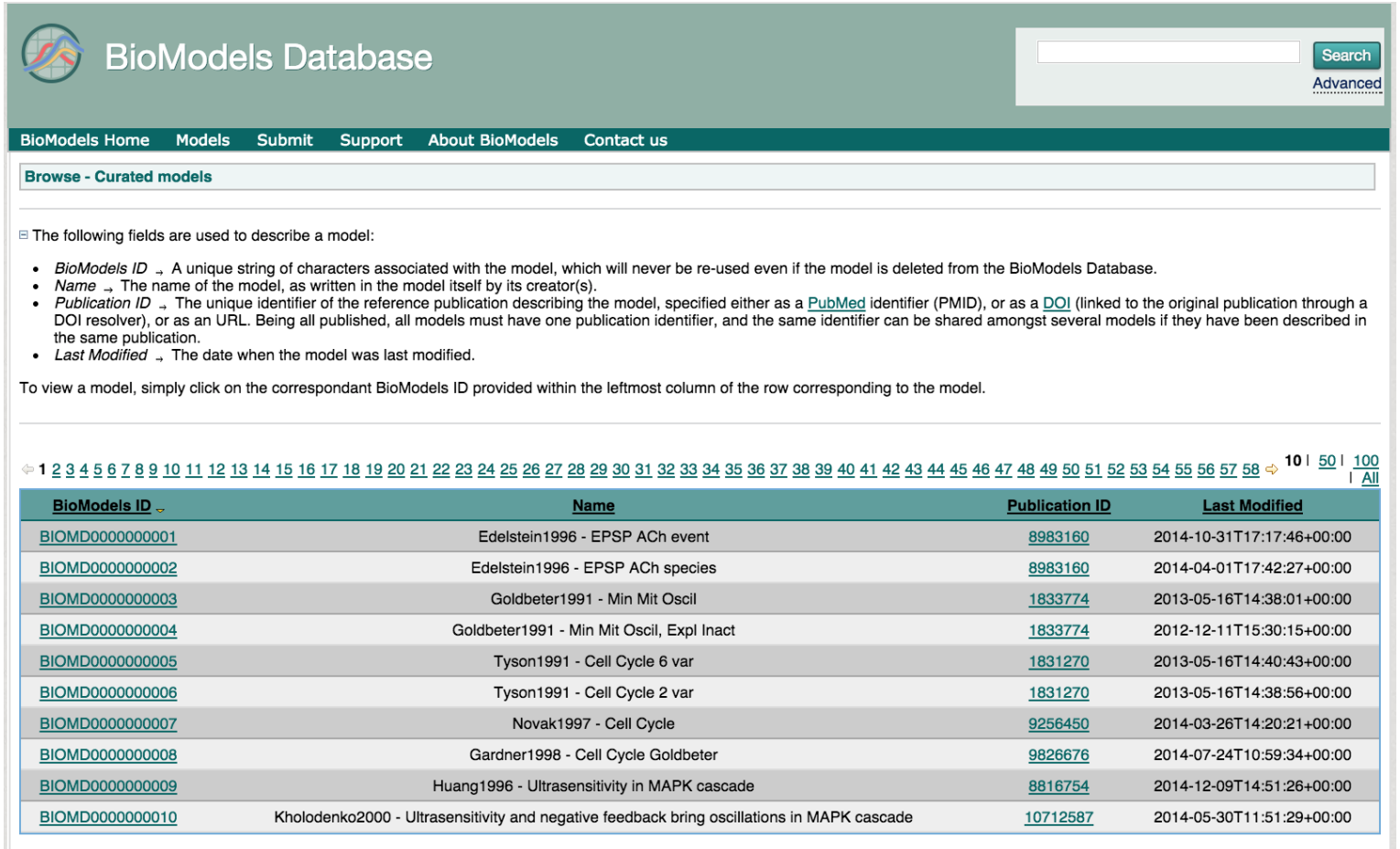


Figure 6B: **The p53-Mdm2 and DNA repair regulatory network** (version 2p - May 19, 1999)

Many formalisms have been developed to interpret them:
ODE, CTMC, Petri nets, Reaction systems, Graph rewriting,
Boolean networks, process algebra...

Model Repositories, e.g. biomodels.net

Systems Biology Markup Language (SBML):
model exchange format (annotations, reactions, events,...)



BioModels Database

Search
Advanced

BioModels Home | Models | Submit | Support | About BioModels | Contact us

Browse - Curated models

The following fields are used to describe a model:

- *BioModels ID* → A unique string of characters associated with the model, which will never be re-used even if the model is deleted from the BioModels Database.
- *Name* → The name of the model, as written in the model itself by its creator(s).
- *Publication ID* → The unique identifier of the reference publication describing the model, specified either as a PubMed identifier (PMID), or as a DOI (linked to the original publication through a DOI resolver), or as an URL. Being all published, all models must have one publication identifier, and the same identifier can be shared amongst several models if they have been described in the same publication.
- *Last Modified* → The date when the model was last modified.

To view a model, simply click on the correspondant BioModels ID provided within the leftmost column of the row corresponding to the model.

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 | 10 | 50 | 100 | All

BioModels ID	Name	Publication ID	Last Modified
BIOMD0000000001	Edelstein1996 - EPSP ACh event	8983160	2014-10-31T17:17:46+00:00
BIOMD0000000002	Edelstein1996 - EPSP ACh species	8983160	2014-04-01T17:42:27+00:00
BIOMD0000000003	Goldbeter1991 - Min Mit Oscil	1833774	2013-05-16T14:38:01+00:00
BIOMD0000000004	Goldbeter1991 - Min Mit Oscil, Expl Inact	1833774	2012-12-11T15:30:15+00:00
BIOMD0000000005	Tyson1991 - Cell Cycle 6 var	1831270	2013-05-16T14:40:43+00:00
BIOMD0000000006	Tyson1991 - Cell Cycle 2 var	1831270	2013-05-16T14:38:56+00:00
BIOMD0000000007	Novak1997 - Cell Cycle	9256450	2014-03-26T14:20:21+00:00
BIOMD0000000008	Gardner1998 - Cell Cycle Goldbeter	9826676	2014-07-24T10:59:34+00:00
BIOMD0000000009	Huang1996 - Ultrasensitivity in MAPK cascade	8816754	2014-12-09T14:51:26+00:00
BIOMD0000000010	Kholodenko2000 - Ultrasensitivity and negative feedback bring oscillations in MAPK cascade	10712587	2014-05-30T11:51:29+00:00

Flat list of models...
Models are annotated separately but not inter-related.

BioModels data

Around 500 "curated" models, 1000 non-curated models.

"Curated" here means that the simulation figures in the reference article could be reproduced by simulating the ODEs.
No consistency condition on their writing in SBML .

Number of species and reactions in some models of interest:

Family	#models	min	max	median
ca_oscil	11	6	44	12
circ	11	24	68	43
mapk	11	20	213	51
cell_cycle	9	7	334	43

From Models to Metamodels

Organize model repositories in model hierarchies[1]:

- define model reductions by a purely structural notion of **subgraph epimorphism (SEPI)** on the **reaction graphs**
- automatically detect model reduction relationships by the existence of SEPI between each pair of models:

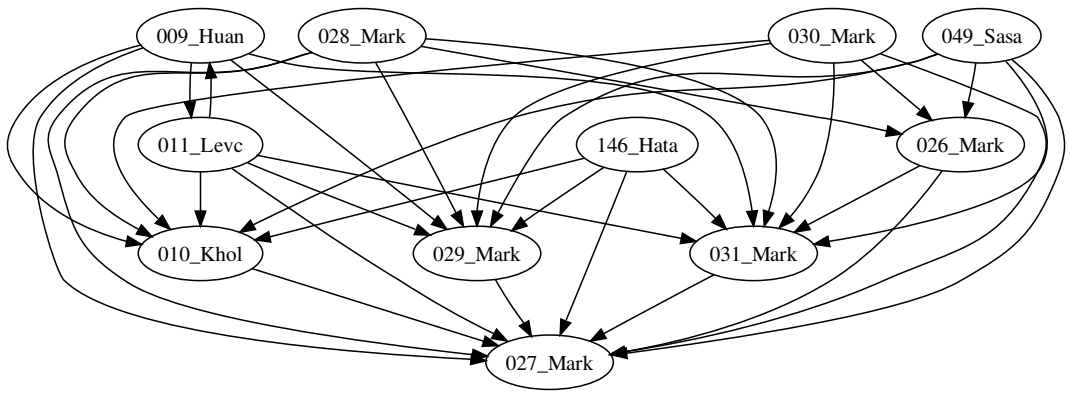


Figure: Hierarchy of MAPK signalling models in biomodels.net, computed by detection of SEPIs

[1] S. Gay F F S. Soliman. A Graphical Method for Reducing and Relating Models in Systems Biology, Bioinformatics 2010

Michaelis-Menten Reduction(s)

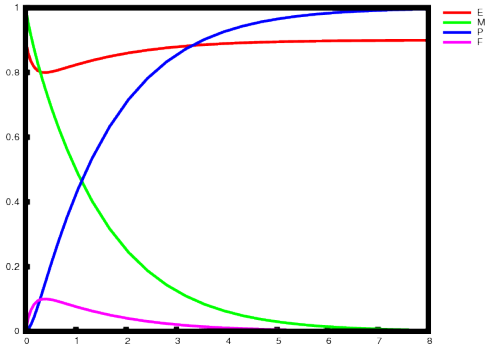
Reaction System = { chemical reactions equipped kinetics} :

Mass-action law kinetics

$k_c \cdot E \cdot S$ for $E + S \to C$

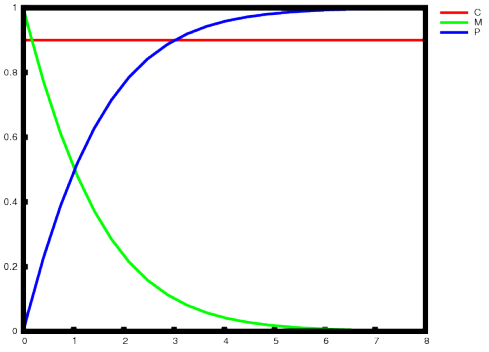
$k_d \cdot C$ for $C \to E + S$

$k_p \cdot C$ for $C \to E + P$



Michaelis-Menten kinetics

$\frac{k_m \cdot E \cdot S}{K_m + S}$ for $E + S \to E + P$



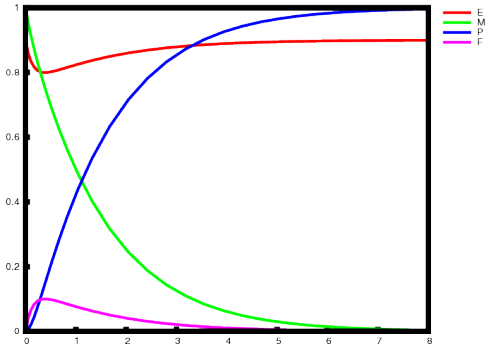
Reduction by projection on a subset of variables (S, P, E)

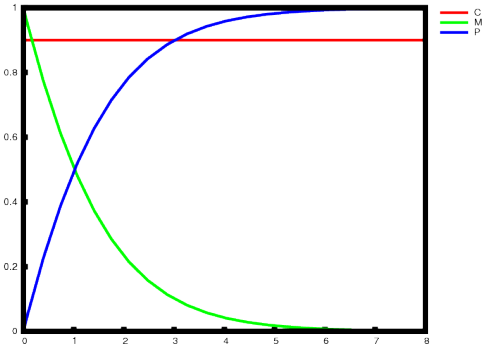
approximated kinetics: **hard to justify and quantify the error**

Michaelis-Menten Reduction(s)

Reaction System = { chemical reactions equipped kinetics} :

Mass-action law kinetics			Michaelis-Menten kinetics
$k_c \cdot E \cdot S$	for	$E + S \to C$	
$k_d \cdot C$	for	$C \to E + S$	$\frac{k_m \cdot E \cdot S}{K_m + S}$ for $E + S \to E + P$
$k_p \cdot C$	for	$C \to E + P$	





Reduction by projection on a subset of variables (S, P, E)
approximated kinetics: **hard to justify and quantify the error**

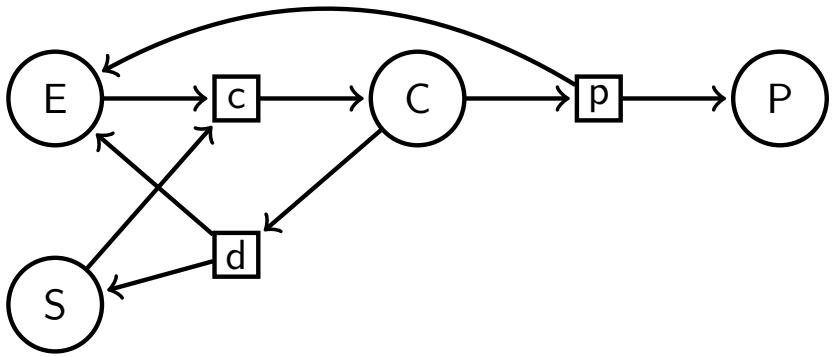
How hell to detect model reductions like that in Biomodels ???

Ignore the Kinetics!

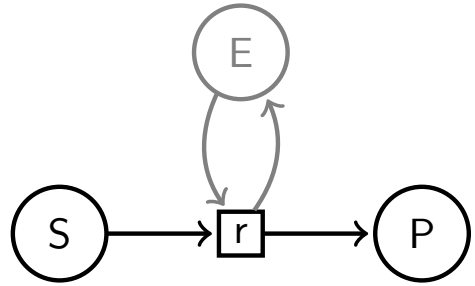
"Si tu ne sais pas résoudre un problème supprime le problème" (proverbe japonais)

Reaction Model = **Reaction Graph** + Kinetics.

Enzymatic mechanism



Michaelis-Menten reduction



Remark: the reduced graph is not isomorphic to a subgraph of the original graph (**no subgraph isomorphism**)

Which graph reduction operations?
Which subgraph morphisms?

Outline

Extracting Reaction Graphs from ODE models

The language of ODEs is low-level (assembler)
The language of reactions is high-level (program)

Many models in BioModels have been transcribed from ODE models with inconsistent reactions.
Extreme case of "curated" model 8: only synthesis reactions with positive and negative terms for the kinetics...

Meaningful reaction systems associated to ODE systems allow for stochastic simulation, Petri net and Boolean interpretations, structural analyses,...

The ODE associated to a reaction system is trivial by CME.
What about the reverse? (`load_ode` in BIOCHAM)
Reaction systems associated to an ODE[2]?

[2] F F, S. Gay, S. Soliman, Inferring Reaction Systems from ODEs, TCS 2014

Well-formed Reaction Systems

A reaction (r, m, p, f) over molecular species $\{x_1, \ldots, x_s\}$ is *well-formed* if the following conditions hold:

1. $f(x_1, \ldots, x_s)$ is a part. diff. function non-negative on $\mathbb{R}^s_+$;
2. $x_i \in r$ if and only if $\partial f / \partial x_i(\vec{x}) > 0$ for some value $\vec{x} \in \mathbb{R}^s_+$;
3. $x_i \in m$ if and only if $\partial f / \partial x_i(\vec{x}) < 0$ for some value $\vec{x} \in \mathbb{R}^s_+$.

Well-formed Reaction Systems

A reaction (r, m, p, f) over molecular species $\{x_1, \ldots, x_s\}$ is *well-formed* if the following conditions hold:

1. $f(x_1, \ldots, x_s)$ is a part. diff. function non-negative on $\mathbb{R}^s_+$;
2. $x_i \in r$ if and only if $\partial f / \partial x_i(\vec{x}) > 0$ for some value $\vec{x} \in \mathbb{R}^s_+$;
3. $x_i \in m$ if and only if $\partial f / \partial x_i(\vec{x}) < 0$ for some value $\vec{x} \in \mathbb{R}^s_+$.

A reaction (r, m, p, f) is *strict* if $f(x_1, \ldots, x_s) = 0$ whenever $x_j = 0$ for any x_j such that $r(x_j) > 0$.

Proposition

The ODE semantics of a well-formed and strict reaction system defines a positive system (i.e. $\mathbb{R}^k_+$ is an invariant set).

$\dot{x} = -k$ is not the ODE of any strict well-formed reaction system, but of the non-strict well-formed reactions $x \xrightarrow{lx} 2x$ and $x \xrightarrow{k+lx} \emptyset$

Influence Graph of a Well-formed Reaction System

The *differential influence graph (DIG)* of a reaction system R is

$$\{x_i \longrightarrow^+ x_j \mid \partial \dot{x_j}/\partial x_i(\vec{x}) > 0 \text{ for some value } \vec{x} \in \mathbb{R}^s_+\}$$
$$\cup \{x \longrightarrow^- x_j \mid \partial \dot{x_j}/\partial x_i(\vec{x}) < 0 \text{ for some value } \vec{x} \in \mathbb{R}^s_+\}$$

The *stoichiometric influence graph (SIG)* of R is

$$\{x \longrightarrow^+ y \mid \text{either } p_i(y) - r_i(y) > 0,\ r_i(x) > 0,$$
$$\text{or } p_i(y) - r_i(y) < 0,\ m_i(x) > 0 \text{ for some reaction } i\}$$
$$\cup \{x \longrightarrow^- y \mid \text{either } p_i(y) - r_i(y) < 0 \text{ and } r_i(x) > 0,$$
$$\text{or } p_i(y) - r_i(y) > 0,\ m_i(x) > 0 \text{ for some reaction } i\}$$

Theorem

For any well-formed reaction system R, $DIG(R) \subset SIG(R)$[3]
and if the $SIG(R)$ has no conflict $DIG(R) = SIG(R)$[4].

---

[3] F F, S. Soliman. From reaction models to influence models and back: a theorem. LNCS. FMSB 2008

[4] F F, S Gay, S. Soliman. Infering reaction systems from ODEs. TCS 2014

Algorithm for Inferring Reactions from ODEs in $O(nt)$

input: ODE system O over variables for molecular concentrations, `partial_has_pos_val` test

1. rewrite O into **additive normal form**
2. compute the set $\mathcal{T}$ of all terms appearing in O
3. let $R := \emptyset$
4. **for each non-decomposable term $t \in \mathcal{T}$ create (r,m,p,f)**
   - 4.1 let $r := \emptyset$, $p := \emptyset$, $m := \emptyset$
   - 4.2 for each x where t occurs with integer coefficient c in $\dot{x}$ in O,
     - 4.2.1 if $c < 0$ then $r(x) := -c$,
     - 4.2.2 if $c > 0$ then $p(x) := c$,
   - 4.3 for each x such that $r(x) = 0$ and `partial_has_pos_val`(t, x),
     - 4.3.1 $r(x) := 1$,
     - 4.3.2 $p(x) := p(x) + 1$,
   - 4.4 for each variable x such that `partial_has_pos_val`$(-t, x)$,
     - 4.4.1 $m(x) := 1$,
   - 4.5 $R := R \cup \{r \ / \ m \xrightarrow{t} p\}$,

output: reaction system R (well-formed whenever possible)

Inferring Ghost Molecules

Variables eliminated by linear invariants lead to factors of the form $c - A - B \ldots$ and non-positive systems (extra condition $X_0 = c - A_0 - B_0$ on the initial state...)

for each expression of the form $k - \Sigma_i x_i$ in O
where k is a numerical constant or a parameter
and x_i are variables,

introduce a new variable z with time derivative $\dot{z} = -\Sigma_i \dot{x}_i$, and functional dependency equation $z = k - \Sigma_i x_i$,

substitute any occurrence of $k - \Sigma_i x_i$ in O by z,

Automatic SBML writing based on ODE semantics

`load_ode` used at Caltech for importing Matlab models in SBML.

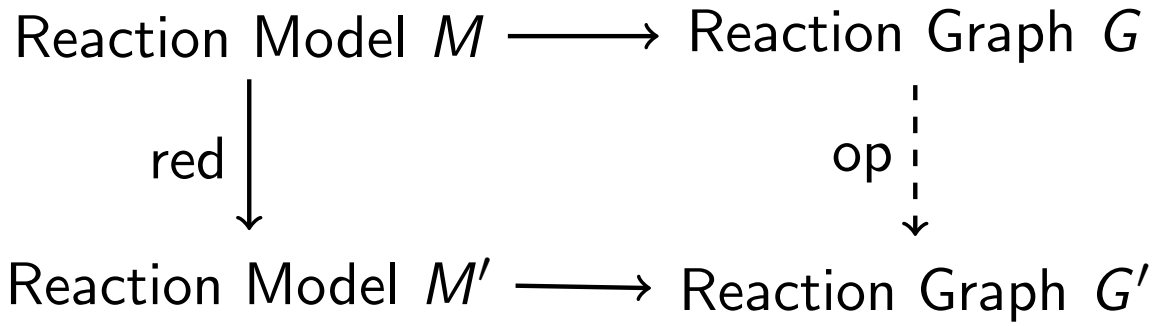
at Montpellier for associating reaction systems to reduced ODEs obtained from tropical equilibrations ?

SBML reaction models can be automatically rewritten in SBML by `export_ODE; load_ODE`

This **decreases the number of inconsistencies from 60% to 30%** in BioModels.net[5].

[5] F F, S Gay, S. Soliman. Infering reaction systems from ODEs. TCS 2014

Graph Operations for Model Reduction



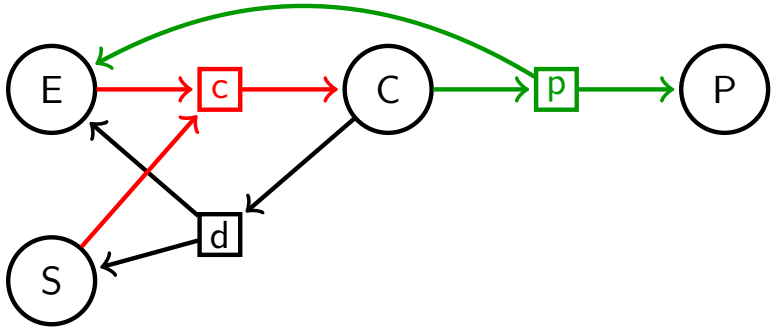
Kinetic considerations	Graph operations
Constant concentration,...	Species deletion
Slow rate,...	Reaction deletion
Proportional concentrations,...	Species merging
Proportional rates,...	Reaction merging

Definition

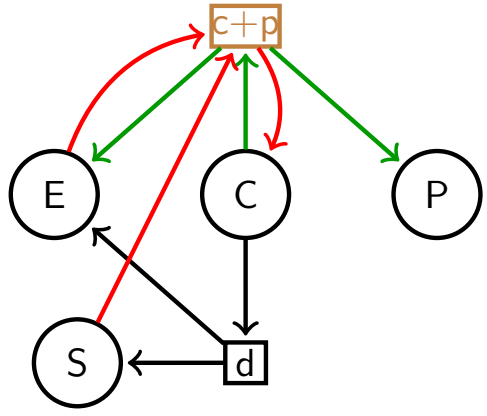
Graphical reduction from G to G' iff chain of delete and merge operations from G to G'

(subgraph isomorphisms correspond to deletions only)

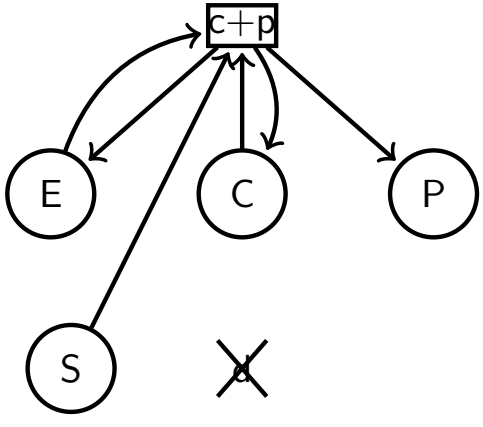
Graph Operations: delete and merge



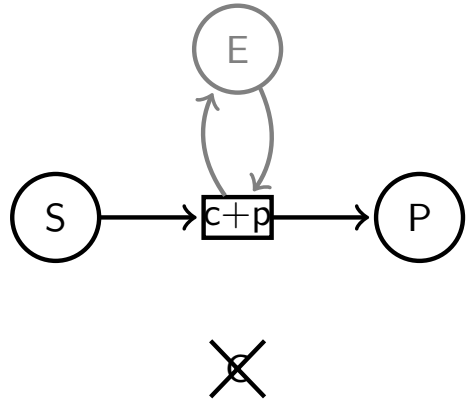
Merge c with p



Delete d



Delete C



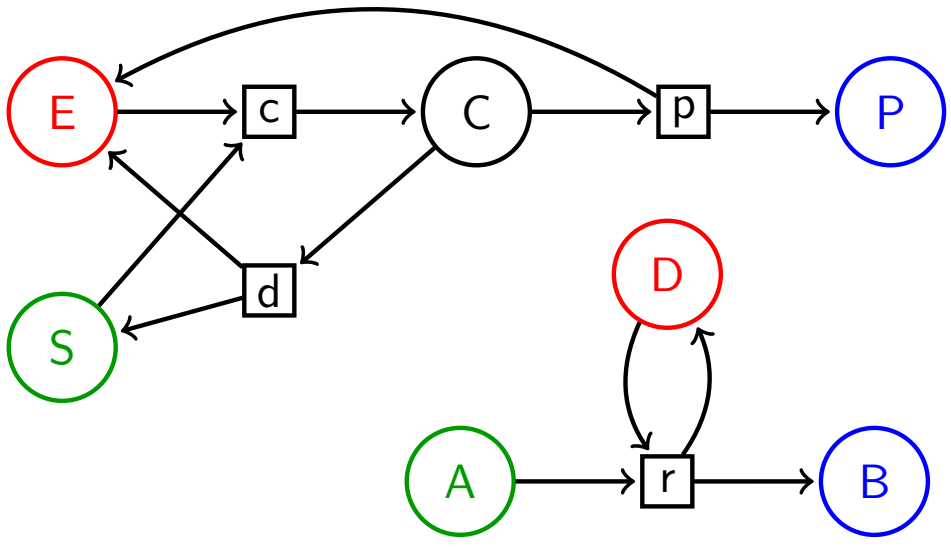
Reduced Michaelis-Menten

Delete/Merge Chains = Subgraph Epimorphisms

Delete/merge chains have many commutation properties: bad for enumerating them.

Flattening the chains' order yields *subgraph epimorphisms*.

	E	$\to$	D
	S	$\to$	A
	P	$\to$	B
merge	c	$\to$	r
merge	p	$\to$	r
delete	d	$\to$	$\bot$
delete	C	$\to$	$\bot$



Subgraph Epimorphisms (SEPI): Formal Definition

Definition

A *subgraph epimorphism* μ is a

- **Function**: from $G = (V, A)$ to $G' = (V', A')$
- **Subgraph**: $\mu : V_0 \subseteq V \longrightarrow V'$ such that:
- **Morphism**: $\forall (x, y) \in A \cap V_0 \times V_0, (\mu(x), \mu(y)) \in A'$
- **Epi**: surjective on the target graph
  - on nodes: $\forall x' \in V', \exists x \in V, \mu(x) = x'$
  - on arcs: $\forall (x', y') \in A', \exists (x, y) \in A, (\mu(x), \mu(y)) = (x', y')$

Theorem

There exists a graphical reduction from G to G' iff there exists a SEPI from G to G'.

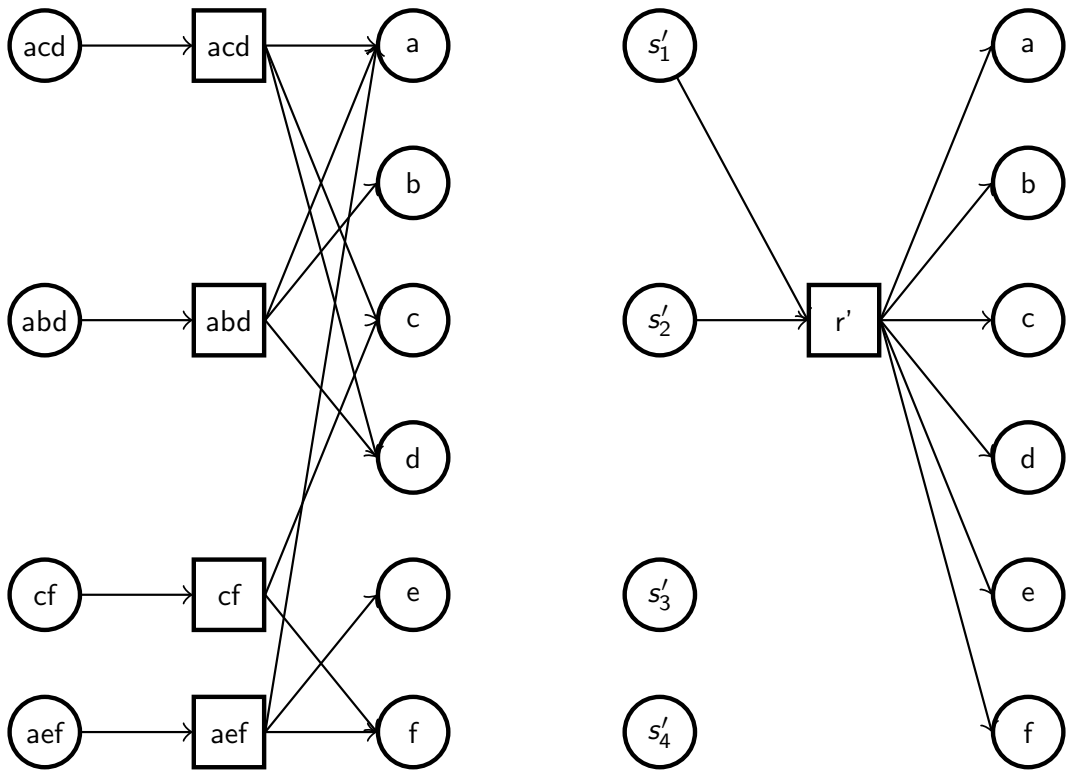
For *reaction graphs*, add $\mu(S_0) \subseteq S'$ and $\mu(R_0) \subseteq R'$ to prevent merging reactions with species.

The SEPI existence problem is NP-Complete

Theorem

The problem of deciding if there is a subgraph epimorphism from G to G' is NP-complete[6]*, even for reaction graphs.*

In the reaction graph case, proved by reduction of k-set-covering:



[6] S. Gay F F T. Martinez S. Soliman C. Solnon, 2014 Discrete Applied Mathematics

Solving SEPI Existence using CP and SAT

CP model characteristics:

- A variable for each node/arc image, $\perp$ for deleted nodes
- Dual model for surjectivity with antecedent variables
- Search on antecedents first, then on direct variables

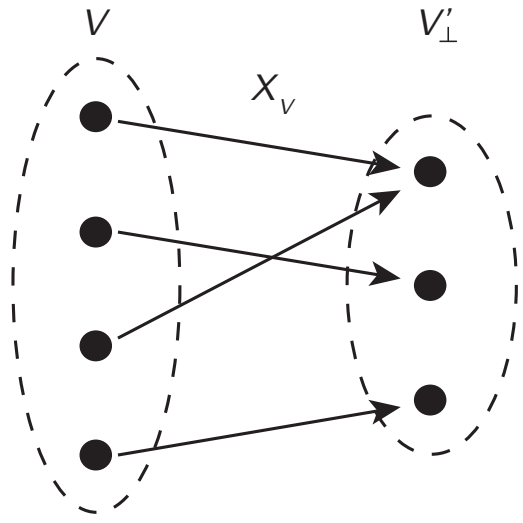
SAT model characteristics:

- Flattened function with order representation (quadratic)
- Surjection with transversal clauses and redundant clauses

CP model for deciding SEPI

Input : $G = (V, A), G' = (V', A')$. Output : yes$(\mu : G \to G')$/no.

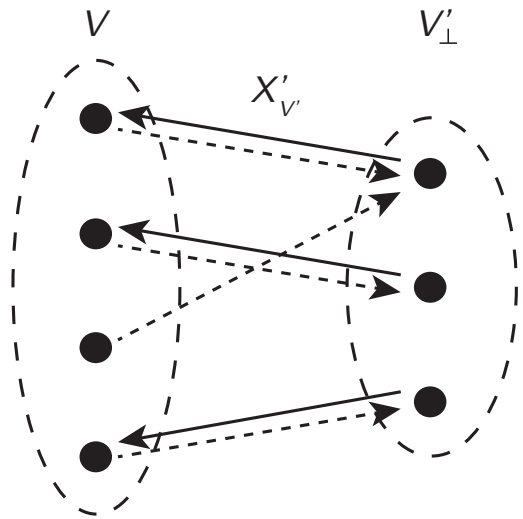
- Morphism variables
  - $\mathbf{X}_v = \mu(v)$ for $v \in V$, with $D(\mathbf{X}_v) = V'_\perp = V' \cup \{\perp\}$.
  - $\mathbf{A}_a = \mu(a)$ for $a \in A$, with $D(\mathbf{A}_a) = A'_\perp = A' \cup (V'^2_\perp - V'^2)$



- Morphism constraints
  - $\forall a \in A, \texttt{element}(\mathbf{A}_a, [\pi_1(a'_1) \dots \pi_1(a'_{|A'_\perp|})], \mathbf{X}_{\pi_1(a)})$
  - $\forall a \in A, \texttt{element}(\mathbf{A}_a, [\pi_2(a'_1) \dots \pi_2(a'_{|A'_\perp|})], \mathbf{X}_{\pi_2(a)})$

CP model for deciding SEPI

- Antecedent variables
  - $\mathbf{X}'_{v'} \in \mu^{-1}(v')$ for $v' \in V'$, with $D(\mathbf{X}'_{v'}) = V$.
  - $\mathbf{A}'_{a'} \in \mu^{-1}(a')$ for $a' \in A'$, with $D(\mathbf{A}'_{a'}) = A$.
- Antecedent constraints, with minimality
  - $\forall v \in V, \forall v' \in V', \mathbf{X}'_{v'} = v \Rightarrow \mathbf{X}_v = v' \wedge \mathbf{X}_v = v' \Rightarrow \mathbf{X}'_{v'} \leq v$
  - $\forall a \in A, \forall a' \in A', \mathbf{A}'_{a'} = a \Rightarrow \mathbf{A}_a = a' \wedge \mathbf{A}_a = a' \Rightarrow \mathbf{A}'_{a'} \leq a$



- Redundant global surjection constraints
  $\texttt{gsurjection}([\mathbf{X}_{v_1} \dots \mathbf{X}_{v_n}], V'), \texttt{gsurjection}([\mathbf{A}_{a_1} \dots \mathbf{A}_{a_k}], A')$
- Strategy: Enumerate $\mathbf{A}'_{a'}$, then $\mathbf{X}'_{x'}$, then $\mathbf{X}_x$ and $\mathbf{A}_a$

SAT model for deciding SEPI

Input : $G = (V, A), G' = (V', A')$. Output : yes$(\mu : G \to G')$/no.

- Variables: Partial surjective function
  - $\forall (v, v') \in V \times (V' \cup \{\bot\})$, $\mathbf{m}_{v,v'} = 1$ iff $m(v) = v'$.
  - $\forall (v, v') \in V \times (V' \cup \{\bot\}), \mathbf{m}^{<}_{v,v'} = 1$ iff $m(v) < v'$.
- Variables: Morphism
  - Non deleted arcs. $\forall (a, a') \in A \times A', \mathbf{m}_{a,a'} = 1$ iff $m(a) = a'$.
  - Deleted arcs. $\forall a \in A, \mathbf{is\_dummy}(\mathbf{m}_a) = 1$ iff $m(a) = \bot$

m	1	2	$\bot$
1	1	0	0
2	0	1	0
3	1	0	0
4	0	0	1

- Clauses: Partial surjective function
  - Left Totality. $\forall v \in V, \bigvee_{v' \in V' \cup \{\bot\}} \mathbf{m}_{v,v'}$
  - Functionality. $\forall (v, v'_j) \in V \times (V' \cup \{\bot\})$,
    $\mathbf{m}_{v,v'_j} \Rightarrow \mathbf{m}^{<}_{v,v'_{j+1}}, \mathbf{m}^{<}_{v,v'_j} \Rightarrow \mathbf{m}^{<}_{v,v'_{j+1}}, \mathbf{m}^{<}_{v,v'_j} \Rightarrow \neg \mathbf{m}_{v,v'_j}$
  - Right Totality. $\forall v' \in V', \bigvee_{v \in V} \mathbf{m}_{v,v'}$

Evaluation on BioModels

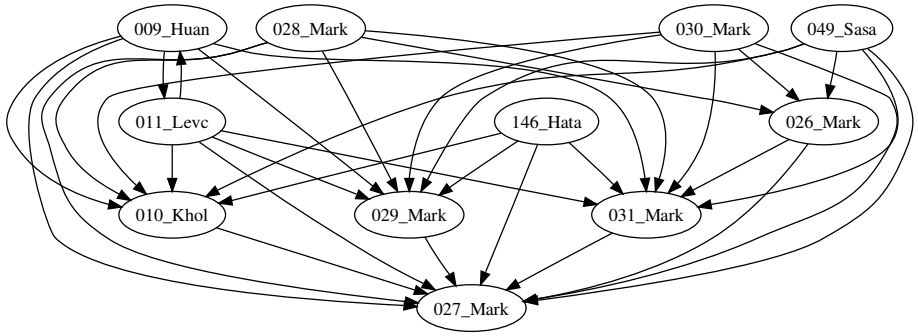
CP model in GNU-Prolog 1.4.4, SAT model with Glucose 2.2.
Evaluation made on 4 classes from the biomodels.net repository.
A comparison result is either SEPI found, no SEPI, or timeout.

Family	Pairs	SEPI			No SEPI			Timeout(30s)		
		CP	SAT	||	CP	SAT	||	CP	SAT	||
ca_oscil	110	38	38	38	72	72	72	0	0	0
circ	110	16	33	33	59	70	70	35	**7**	7
mapk	110	37	37	41	59	62	62	14	11	**7**
cell_cycle	72	9	8	10	42	49	49	21	15	**13**

Between different clusters, only 9% of SEPI detections mostly from some small to large models.

Work to do: search **small frequent SEPI motifs** in BioModels!
(open problem for protein networks)

SEPI Partial Order



$(\mathcal{G}, \leq_{SEPI})$ is a partial order.
Let $G_{\downarrow} = \{H \mid H \leq_{SEPI} G\}$, $G_{\uparrow} = \{H \mid G \leq_{SEPI} H\}$.

A **maximal element** of $G_{\downarrow} \cap G'_{\downarrow}$ captures **common structure** of G and G', by removing a minimal amount of specific structure.
Their set is written $G \sqcap G'$, we call them **SEPI "intersections"**.

A **minimal element** of $G_{\uparrow} \cap G'_{\uparrow}$ captures **specific structure** of G and G', by sharing of maximal amount of common structure.
Their set is written $G \sqcup G'$, we call them **SEPI "unions"**.

Properties of SEPI union, intersection

$(\mathcal{G}, \leq_{SEPI})$ is a not a lattice.

Theorem ("Intersections" properties)

No greatest element: $\exists G, G', |G \sqcap G'| \geq 1$.
Maximals can have $\neq$ sizes: $\exists G, G', \exists I, I' \in G \sqcap G', |I| \neq |I'|$.
Size bound: $\forall I \in G \sqcap G', |I| \leq min(|G|, |G'|)$.

Theorem ("Unions" properties)

No smallest element: $\exists G, G', |G \sqcup G'| \geq 1$.
Minimals can have $\neq$ sizes: $\exists G, G', \exists U, U' \in G \sqcup G', |U| \neq |U'|$.
Size bound: $\forall U \in G \sqcup G', |U| \leq (|G| + 1) \cdot (|G'| + 1)$.

$(\mathcal{G}, \leq_{SEPI})$ is a not a well-quasi order (infinite SEPI-antichains).
Fundamental difference with graph minors (e.g. cannot be used to detect planarity by reduction to specific graphs)

SEPI Distance

Consider $\mathcal{G}$ with one-step merge/delete transitions $\rightarrow_{md}$.
The walks in $(\mathcal{G}, \rightarrow_{md})$ define a natural distance:

Definition

Let $d_{SEPI}(G, G')$ be the length of a shortest walk from G to G'.

Theorem

There is always a shortest walk of the form $G \rightarrow_{md}^{*} I \leftarrow_{md}^{*} G'$, *with* $I \in G \sqcap G'$.

Corollary

$\forall I \in G \sqcap G'$ *of maximal size,* $d_{SEPI}(G, G') = |G| - |I| + |G'| - |I|$

These results are not true for $G \sqcup G'$.

Prototypes

A CP model (in GNU-Prolog) has been attempted for SEPI intersection, answers for very few instances.

SAT models have been evaluated for SEPI intersection and union, answers for 25% of the instances.

How hell to get rid of those **too many SEPIs ?**

Don't Ignore the Kinetics !

"Impossible n'est pas français" (proverbe français)

Michaelis-Menten Reduction by Tropical Equilibration:

- Let $x_1 = [S]$, $x_2 = [ES]$, $x_3 = [E]$, $x_4 = [P]$, we have
  $x_1' = -k_1 x_1 x_3 + k_{-1} x_2$
  $x_2' = k_1 x_1 x_3 - (k_{-1} + k_2) x_2$
  $x_3' = -k_1 x_1 x_3 + (k_{-1} + k_2) x_2$
  $x_4' = k_2 x_2.$
  $x_2 + x_3 = e_0$
  $x_1 + x_2 + x_4 = s_0$
- **Rescaling**, $x_i = \bar{x}_i \epsilon^{a_i}$,
  $k_1 = \bar{k}_1 \epsilon^{\gamma_1}$, $k_{-1} = \bar{k}_{-1} \epsilon^{\gamma_{-1}}$, $e_0 = \bar{e}_0 \epsilon^{\gamma_e}$, $s_0 = \bar{s}_0 \epsilon^{\gamma_s}$.
- **Tropical equilibration equations** over **(min,+) semiring** :
  $\gamma_1 + a_1 + a_3 = \gamma_{-1} + a_2$
  $\gamma_1 + a_1 + a_3 = \min(\gamma_{-1}, \gamma_2) + a_2$
  $\gamma_1 + a_1 + a_3 = \gamma_2 + a_2$
  $\min(a_2, a_3) = \gamma_e$
  $\min(a_1, a_2, a_4) = \gamma_s$
  The solutions over the integers provide reductions that generalize QSSA and QE reductions.

Tropical Equilibration Algorithm?

- **Inputs**:
  - Biochemical reaction system with mass action kinetics,
  - small ϵ,
  - conservation laws (can be computed as Petri net Place-invariant by a CSP [Soliman 2012]).

Tropical Equilibration Algorithm?

- **Inputs**:
  - Biochemical reaction system with mass action kinetics,
  - small ϵ,
  - conservation laws (can be computed as Petri net Place-invariant by a CSP [Soliman 2012]).

- **Output**:
  - Rescaling of the variables s.t.
    - the dominating positive and negative term in each ODE equilibrate (are of the same degree in ϵ),
    - the conservation laws hold.

Tropical Equilibration as a Constraint Satisfaction Problem

- **Variables**: $a_i \in \mathbb{Z}$ (for $x_i = \epsilon^{a_i} \bar{x}_i$)

[7] S. Soliman, F F, O. Radulescu. A constraint solving approach to model reduction by tropical equilibration. Algorithms for Molecular Biology, 9(24), 2014

Tropical Equilibration as a Constraint Satisfaction Problem

- **Variables**: $a_i \in \mathbb{Z}$ (for $x_i = \epsilon^{a_i} \bar{x}_i$)
- **Constraints**:
  For each species
  min(PositiveMonomialDegrees, M)
  min(NegativeMonomialDegrees, M)
  For each conservation law with K= $\mathrm{round}(\log(c_i)/\log(\epsilon))$.
  min(ConservationLawDegrees, K)

[7] S. Soliman, F F, O. Radulescu. A constraint solving approach to model reduction by tropical equilibration. Algorithms for Molecular Biology, 9(24), 2014

Tropical Equilibration as a Constraint Satisfaction Problem

- **Variables**: $a_i \in \mathbb{Z}$ (for $x_i = \epsilon^{a_i} \bar{x}_i$)
- **Constraints**:
  For each species
  min(PositiveMonomialDegrees, M)
  min(NegativeMonomialDegrees, M)
  For each conservation law with K= $\mathrm{round}(\log(c_i)/\log(\epsilon))$.
  min(ConservationLawDegrees, K)
- **Search**: Domain bisection with iterative domain doubling from $[-2, 2]$ (i.e. in $[10^{-2}, 10^2]$) to $[-128, 128]$.

[7] S. Soliman, F F, O. Radulescu. A constraint solving approach to model reduction by tropical equilibration. Algorithms for Molecular Biology, 9(24), 2014

Tropical Equilibration as a Constraint Satisfaction Problem

- **Variables**: $a_i \in \mathbb{Z}$ (for $x_i = \epsilon^{a_i} \bar{x}_i$)
- **Constraints**:
  For each species
  min(PositiveMonomialDegrees, M)
  min(NegativeMonomialDegrees, M)
  For each conservation law with K= $\mathrm{round}(\log(c_i)/\log(\epsilon))$.
  min(ConservationLawDegrees, K)
- **Search**: Domain bisection with iterative domain doubling from $[-2, 2]$ (i.e. in $[10^{-2}, 10^2]$) to $[-128, 128]$.

Use of a solver handling $\mathbb{Z}$ like SWI-Prolog `clpfd` library[7].

[7] S. Soliman, F F, O. Radulescu. A constraint solving approach to model reduction by tropical equilibration. Algorithms for Molecular Biology, 9(24), 2014

Reified Cardinality Constraint on Minimum

The minimum value must be reached only once (steady state) but at least twice

Constraint min(L, M, N) true if M is smaller than each element of L and equal to N elements of that list.

Reified Cardinality Constraint on Minimum

The minimum value must be reached only once (steady state) but at least twice

Constraint min(L, M, N) true if M is smaller than each element of L and equal to N elements of that list.

Domain filtering between L, M and N by reified constraints:

```
min([], _, 0).
min([H | T], M, C) :-
        M=<H, B <==> M=H, C=B+CC,
        min(T, M, CC).

min(M, L) :-  C>=1, min(M, L, C).
```

Reified Cardinality Constraint on Minimum

The minimum value must be reached only once (steady state) but at least twice

Constraint min(L, M, N) true if M is smaller than each element of L and equal to N elements of that list.

Domain filtering between L, M and N by reified constraints:

```
min([], _, 0).
min([H | T], M, C) :-
        M=<H, B <==> M=H, C=B+CC,
        min(T, M, CC).

min(M, L) :-  C>=1, min(M, L, C).
```

Constraint propagation could be further improved using minimum and min_n **global constraints** [Beldiceanu et al 2005]

Results on Biomodels.net

55 out of 436 curated models have non-trivial purely polynomial kinetics.

Results on Biomodels.net

55 out of 436 curated models have non-trivial purely polynomial kinetics.

$\epsilon = 0.1$

Found	# models	# var (avg/min/max)	Time (avg/min/max)
yes	23	17.348/3/ 86	0.486/0.004/2.803
no	32	17.812/1/194	0.099/0.000/1.934

Results on Biomodels.net

55 out of 436 curated models have non-trivial purely polynomial kinetics.

$\epsilon = 0.1$

Found	# models	# var (avg/min/max)	Time (avg/min/max)
yes	23	17.348/3/ 86	0.486/0.004/2.803
no	32	17.812/1/194	0.099/0.000/1.934

18 of 23 models with equilibrations appear to have an infinite number of equilibrations (e.g. linear relations between variables)

Model	# equilibrations	Total time (s)
BIOMD0000000002	36	109
BIOMD0000000122	45	291
BIOMD0000000156	7	0.008
BIOMD0000000229	7	0.7
BIOMD0000000413	29	3.3

References


François Fages, Steven Gay, and Sylvain Soliman.
Inferring reaction systems from ordinary differential equations.
Theoretical Computer Science, September 2014.


Steven Gay, François Fages, Thierry Martinez, Sylvain Soliman, and Christine Solnon.
On the subgraph epimorphism problem.
Discrete Applied Mathematics, 162:214–228, January 2014.


Steven Gay, Sylvain Soliman, and François Fages.
A graphical method for reducing and relating models in systems biology.
Bioinformatics, 26(18):i575–i581, 2010.
special issue ECCB'10.


Sylvain Soliman, François Fages, and Ovidiu Radulescu.
A constraint solving approach to model reduction by tropical equilibration.
Algorithms for Molecular Biology, 9(24), December 2014.